Kidney International, Vol. 41 (1992), pp. 645–649

Peptide recognition motifs involved in the binding of integrins to their ligands

MARTIN J. HUMPHRIES

Department of Biochemistry and Molecular Biology, University of Manchester, Manchester, England, United Kingdom

For both cell-cell and cell-matrix interactions, the net adhesive event is now known to be the product of a multitude of individual binding interactions mediated by a variety of adhesive ligands and counter-receptors. Thus, complex spatial arrangements of membrane-intercalated molecules at the cell surface interact specifically with complementary surfaces of other cells or with the pre-assembled macromolecular complexes of glycoproteins and proteoglycans that make up most extracellular matrices. Despite this complexity, adhesion is a specific process that is capable, for example, of promoting directional cell migration and modulating tissue-specific gene expression. To understand how the phenotypic effects of adhesion are generated, studies of the structures and functions of the molecules mediating adhesive recognition have been performed, and these have, in recent years, provided significant insights into the mechanisms involved.

Identification of RGD-containing cell-binding sequences in adhesive macromolecules

Following the identification and purification of the major glycoprotein components of extracellular matrices, rapid progress has been made in mapping the biological activities of these molecules to particular protein domains. Thus, for example, regions of fibronectin, laminin and the interstitial collagens that mediate cell-binding have been isolated as proteolytic fragments from their parent molecules. Most rapid progress in identifying the minimal cell adhesive segments of these molecules has been made with fibronectin, largely because it is extremely amenable to dissection with enzymes or chemicals under conditions that result in retention of biological activity. By examining the adhesive activity of progressively smaller fragments of fibronectin, one key active site in the centre of the molecule was eventually reproduced in synthetic peptide form, the minimal sequence of which was a tetrapeptide RGDS [1] (Fig. 1). RGDS-containing peptides were found to support cell adhesion directly and to act as competitive inhibitors of fibronectin-mediated adhesion.

The surprising finding that such a short peptide could mediate such a complex process as cell adhesion has nonetheless been corroborated by a variety of further studies. First, although in structure-function studies a number of amino acids other than S were found to be tolerated in the fourth position of the tetrapeptide, peptides containing substitutions in either of the first three amino acids of the tetrapeptide lacked functional activity, indicating a precise requirement for RGD and therefore that the biological effects of the peptides were specific [2]. Second, deletion of RGDS or mutation of RGDS to RGES in fusion proteins expressed from cDNA clones spanning the central cell-binding domain of fibronectin resulted in almost complete loss of adhesive activity [3]. Third, the cloning and sequencing of extracellular matrix molecules has revealed that most contain RGD-type motifs. Furthermore, evidence has been presented from the use of generic RGD peptide inhibitors to block the adhesive function of RGD-containing molecules that in most cases these sites are functional [4]. Although non-RGD cell-binding sites exist, it is clear that RGD represents a common adhesive recognition motif.

Identification of adhesion receptors

The identification of the cell surface molecules acting as functional adhesion receptors for extracellular matrix molecules initially proved difficult owing to the relatively low affinity of the interactions involved. For fibronectin, for example, direct binding studies revealed an affinity constant of 0.8 μM for the intact molecule which is borderline for obtaining successful binding in affinity chromatography experiments [5]. Somewhat fortuitously, fragments of fibronectin containing the central cell-binding domain of fibronectin were found to bind with slightly higher affinity to intact cells and it proved possible to use these fragments for affinity isolation. In this way, the central cell-binding domain receptor was shown to be a dimer of molecular weight 100 to 140 K that bound to affinity matrices in an RGDS-sensitive manner [6]. After cloning and sequencing of both subunits, it became clear that this fibronectin receptor was one member of a large family of surface molecules which are now called the integrins.

Integrins are composed of non-covalently-bound α and β subunits, of which there are currently 12 and 7 characterized examples, respectively, in mammalian cells [4]. These 19 different subunits have been shown to assemble into 17 different dimers, each of which exhibits a different ligand-binding specificity [4]. As shown in Figure 2, although there are examples of integrin α subunits that can associate with more than one β subunit, the most convenient method of classification of a particular dimer is according to its β subunit. As a generalization, β1 integrins appear to be principal mediators of cell-extracellular matrix adhesion, with fibronectin, laminin and

© 1992 by the International Society of Nephrology

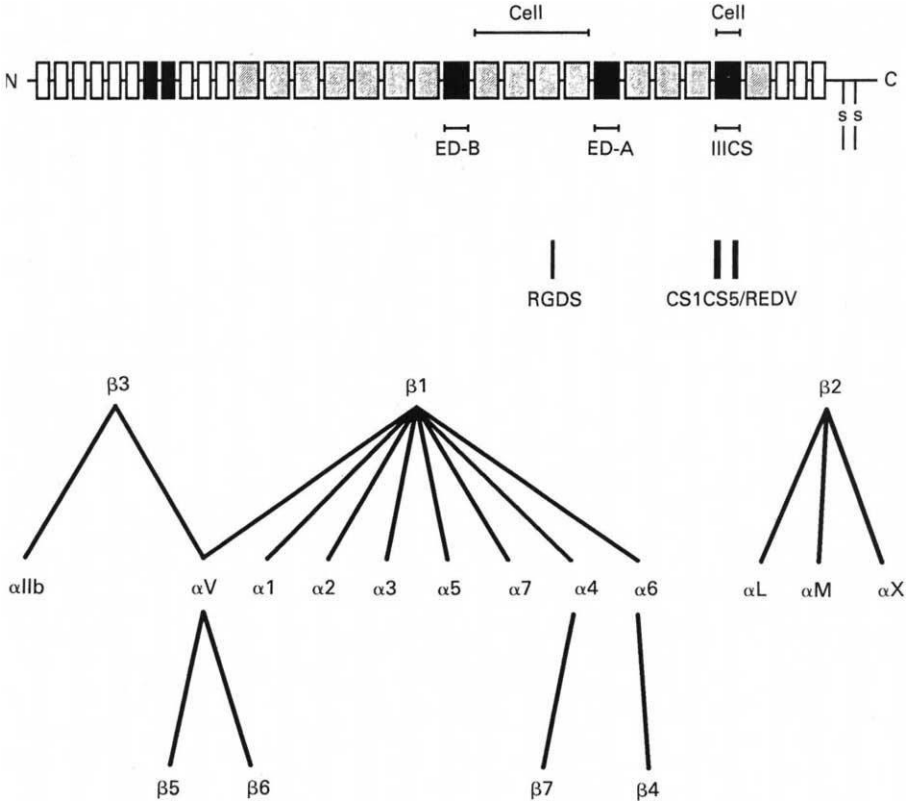

Fig. 1. *Schematic structure of the fibronectin subunit.* Fibronectin is a dimer linked by disulfide bonds close to the COOH-termini of its subunits. Each subunit has a similar structure and is composed of a number of homologous repeating modules (shown by rectangles). These modules can be classified into three types (shown by open, hatched and cross-hatched shading). Fibronectin also contains three sites for alternative mRNA splicing (shown by solid shading). The two cell-binding domains within fibronectin are located between the two ED splice segments (the central cell-binding domain) and in the third splice segment (the IIICS). The central cell-binding domain contains the RGDS tetrapeptide as one key cell-binding site, while the IIICS contains at least two sites, CS1 which does not have an RGD motif and CS5 which does (RGDV in rat and bovine fibronectins and REDV in human fibronectin). Quantitatively, CS1 dominates the activity of the IIICS.

Fig. 2. *Known mammalian integrin subunits and dimer combinations.* According to current nomneclature, 12 α subunits and 7 β subunits have been identified. The lines linking particular α and β subunits indicate known dimeric combinations. A detailed description of the ligand-binding specificities of particular integrins can be found in reference 4.

collagens as their major ligands, $\beta 2$ integrins are involved in leukocyte cell-cell adhesion and $\beta 3$ integrins are related in function to the $\beta 1$ molecules, except that they appear to bind to a different set of ligands including vitronectin, fibrinogen and von Willebrand factor. The $\beta 4$ subunit is currently unique among integrins, both in terms of its structure and proposed function. In contrast to all other β subunits, $\beta 4$ has an extended cytoplasmic domain, and this has been suggested to mediate interactions with intermediate filaments in epithelial hemidesmosome structures [7]. At present, the $\beta 5$, $\beta 6$ and $\beta 7$ subunits are only known to associate with single α subunits, and the quantitative importance of each of these molecules is difficult to assess.

Structurally, all integrin α subunits are homologous to each other, as are the β subunits; the pertinent features of each type of subunit are shown in Figure 3. In general, both types of subunit are composed of large extracellular domains, putative transmembrane regions and short cytoplasmic domains. Some α subunits are proteolytically processed at an extracellular site close to the region of membrane intercalation into a disulfide-linked heavy and light chain, others possess an inserted (or I) domain close to the NH_2-terminus. All α subunits, however, possess a series of seven repeating polypeptide units, the last three or four of which contain putative divalent cation-binding sites homologous to the EF-hand structure identified in parvalbumin and calmodulin. The possible significance of these sites is discussed in more detail below. Integrin β subunits, with the exception of $\beta 4$, have few distinguishing features, although all possess a series of extracellular, cysteine-rich repeats which appear to explain the anomalous elecrophoretic behavior of these subunits.

Non-RGD integrin-binding ligand active sites

Although RGD is now well established as an integrin recognition sequence in many adhesive proteins, for others, domains that lack RGD sequences have been shown to exhibit cell adhesion-promoting activity and for some integrins, ligand binding has been shown to be RGD-insensitive. Thus, it is likely that other mechanisms of integrin-ligand binding exist. There are currently two examples in which integrin-binding sites have been localized within adhesion proteins to non-RGD sequences: these are the IIICS domain of fibronectin and the γ-chain of fibrinogen.

The IIICS region of fibronectin is a 120-amino acid-long, cell type-specific adhesive domain recognized by neural crest derivatives and mononuclear leukocytes. Two active sites have

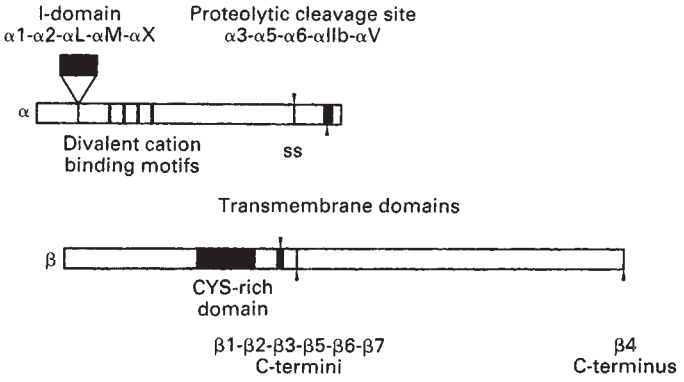

Fig. 3. *Integrin subunit structures.* With the exception of $\beta 4$, all integrin subunits have short cytoplasmic and long extracellular domains. Sites of proteolytic processing, I-domain insertion and intramolecular disulfide bonding are shown.

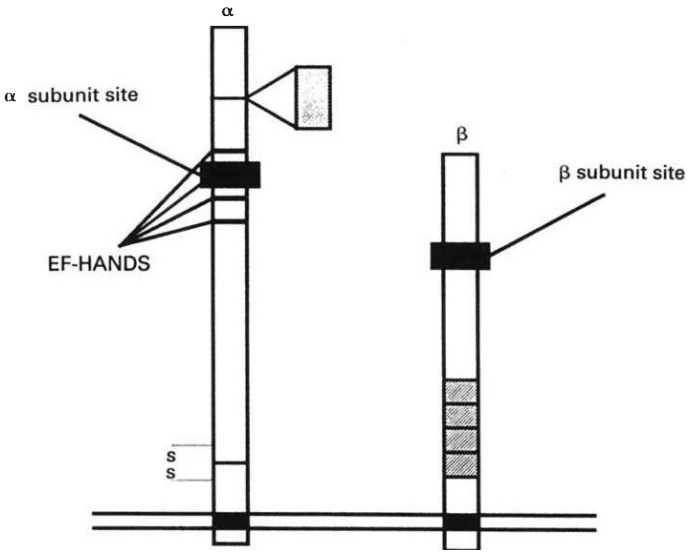

Fig. 4. *Integrin active sites.* Currently, two regions on integrin α and β subunits have been proposed as ligand-binding sites. As discussed in the text, these are the second EF-hand motif in αIIb and αV and a region in the NH_2-terminal half of $\beta 3$.

been pinpointed using synthetic peptides within the IIICS; these are represented by the 25- and 20-mer peptides CS1 and CS5 (Fig. 1). Whereas the minimal active sequence within CS5 is a tetrapeptide incorporating the RGD recognition motif, the CS1 peptide, which has a dominant contribution to the adhesive activity of the IIICS, does not possess an RGD sequence [8, 9]. Antibody-blocking and affinity chromatography experiments have identified the receptor for both the CS1 and CS5 active sites as the integrin $\alpha_4\beta_1$, thus demonstrating in the case of CS1 that not all integrin-ligand interactions involve RGD [10, 11]. In recent studies, the minimal active sequence within CS1 has in fact been narrowed down to a second tripeptide, LDV [12]. Interestingly, data suggest that RGDX and LDV may be functionally equivalent; substitution of a glutamic acid residue for the aspartate residue in either peptide causes an almost complete loss of activity (demonstrating that both sequences have an absolute dependence on their D residue) and both peptides are able to competitively inhibit each other's function [13].

The second well-characterized, non-RGD integrin-binding sequence is found at the extreme COOH-terminus of the γ-chain of fibrinogen. This region, which interacts directly with the integrin $\alpha IIb\beta_3$ (also known as platelet glycoprotein IIbIIIa), can be represented by the dodecapeptide HHLGGAKQAGDV [14]. As for RGD and LDV, HHLGGAKQAGDV activity is also dependent on the penultimate D residue for activity and exhibits the same competitive inhibition of function with the active RGD site in the α chain of fibrinogen [15].

The equivalence of function between HHLGGAKQAGDV and RGD and LDV and RGD suggests that each of these sequences may be a common recognition signal employed by adhesion proteins for integrin binding. Although no data are available for HHLGGAKQAGDV, in preliminary studies, three adhesion protein peptides that are homologous, but not identical, to the LDV-containing peptide CS1 have been examined for adhesive activity. The sequences were taken from the HepII proteoglycan-binding domain of fibronectin (a region which is adjacent to the IIICS and known to be recognized by $\alpha_4\beta_1$) and from two other proteins, the cytokine-inducible endothelial cell surface protein VCAM-1 (also known to be recognized by $\alpha_4\beta_1$) and the lymphocyte surface antigen CD45. VCAM-1 has been implicated in leukocyte trafficking across endothelial monolayers, while the function of CD45 is unknown despite the strong correlation between expression of its various isoforms and memory cell phenotype. In all three cases, using either cell attachment or spreading assays, peptides containing these sequences were found to support adhesion of either lymphoblastic leukemia or melanoma cells (Markarem, Mould and Humphries, unpublished observations). In addition, for all these peptides, adhesion was mediated by $\alpha_4\beta_1$. Thus, preliminary indications are that LDV, like RGD, may be a common integrin-binding motif.

Hypotheses to explain integrin-ligand binding specificity

Despite the inherent complexities of cell adhesion outlined in the introduction to this overview, the findings described above indicate that there may be at least be some common molecular basis to many of the intermolecular interactions that constitute an adhesive event. First, the majority of data to date indicate that adhesive ligands possess and use related aspartate-containing peptide motifs as cell-binding sites. Second, the cell surface receptors recognizing these sequences all belong to the integrin gene family. Thus, in the simplest case, it is conceivable that adhesive specificity may be determined by the recognition of particular conformations of ligand-active site peptides by structurally homologous binding sites in integrin dimers. To test this hypothesis, it is first necessary to identify the sites on integrin receptors that interact with these sequences.

Information on the ligand-binding sites in the $\beta 3$ integrins, $\alpha V\beta_3$ and $\alpha IIb\beta_3$, has been obtained initially from chemical cross-linking studies (Fig. 4). Using radiolabelled peptides, RGD- and HHLGGAKQAGDV-binding sites have been found on the αIIb subunit, while a second RGD-binding site has been localized to a region of the $\beta 3$ subunit [16–19]. In the case of the HHLGGAKQAGDV peptide, the cross-linking site was found to be the second of the four EF-hand-type cation-binding repeats in αIIb. Interestingly, the RGD cross-linking site on $\beta 3$

also shows some homology to the EF-hand consensus sequence, suggesting that these two types of peptide may bind close to or at sites of divalent cation sequestration in integrins. Although the binding sites for LDV-containing peptides have not yet been identified, as discussed above, information from studies of the competitiveness of peptide inhibition of cell adhesion suggest that RGD and LDV peptides may share a receptor-binding site on $\alpha_4\beta_1$ [13]. The potential importance of the polypeptide around the RGD cross-linking site in $\beta 3$ is enhanced by the finding that a mutation in this region of the molecule can result in the Cam variant of Glanzmann's thrombasthenia [20]. This condition is a bleeding disorder characterized by aberrant fibrinogen-dependent platelet aggregation, and is manifested by abnormal cation-dependent conformational changes in $\alpha IIb\beta_3$. Similarly, a number of the mutations causing leukocyte adhesion deficiency have been mapped to a homologous region of the $\beta 2$ subunit [21]. In recent studies, a synthetic peptide taken from the HHLGGAKQAGDV cross-linking site on αIIb has been shown to block both platelet aggregation and binding of fibrinogen to $\alpha IIb\beta_3$, thus providing direct evidence that this site is of functional relevance [22].

The presence of cation-binding sites in integrin subunits appears to explain the dependency of integrin-ligand binding and cell adhesion on divalent cations. Interestingly, however, examination of the sequences of the EF-hand repeats in integrin α subunits reveals an interesting anomaly: one of the six residues that is normally involved in coordinating the bound cation is absent in all integrin EF-hands. Usually this residue is either an aspartate or a glutamate. However, in integrins it is replaced by a small hydrophobic amino acid. This has led to the intriguing hypothesis that the aspartate residue shown to be critical for the functioning of the ligand motifs, RGD, LDV and HHLGGAKQAGDV, may in fact function by providing its side-chain for direct coordination of cation [20, 23]. Thus, adhesive specificity may be generated by the interaction of specific conformations of aspartate-containing active sites in extracellular matrix proteins with particular "defective" EF-hand structures in integrin subunits. This hypothesis could explain the findings that multiple integrins can bind to the same site in the same extracellular matrix molecule and that multiple ligands can be bound by the same integrin [4]. Alternatively, the interaction of D-containing peptides with EF-hand structures may serve a different function in activation of integrins. There is now considerable evidence that integrins can exist in both low and high avidity states and that the transient activation of the receptors is of significant physiological relevance [24]. It is possible that initial binding of D-containing peptide sites to EF-hands, rather than being the principal adhesive interaction, may instead elicit conformational changes in integrins that enhance and stabilize further ligand interactions. This hypothesis is supported by the finding that soluble RGD peptides are able to induce post-receptor binding events in $\alpha IIb\beta_3$ [25]. Clearly, in addition to further peptide-based studies of ligand and receptor active sites in adhesion molecules, one major area of future investigation will be studies aimed at obtaining a three-dimensional description of an integrin-ligand interaction in order to assess the relative functional roles of aspartate-containing peptide motifs and other cell-binding sequences.

Acknowledgments

Grateful acknowledgment is made to The Wellcome Trust and the Wigan and District Cancer Research Committee for financial support.

Reprint requests to Dr. Martin J. Humphries, Department of Biochemistry and Molecular Biology, University of Manchester, Stopford Building, Oxford Road, Manchester, M13 9PT, England, United Kingdom.

References

1. Pierschbacher MD, Ruoslahti E: The cell attachment activity of fibronectin can be duplicated by small synthetic fragments of the molecule. *Nature* 309:30–33, 1984
2. Pierschbacher MD, Ruoslahti E: Influence of stereochemistry of the sequence Arg-Gly-Asp-Xaa on binding specificity in cell adhesion. *J Biol Chem* 262:17294–17298, 1987
3. Obara M, Kang MS, Yamada KM: Site-directed mutagenesis of the cell-binding domain of human fibronectin: Separable, synergistic sites mediate adhesive function. *Cell* 53:649–657, 1988
4. Humphries MJ: The molecular basis and specificity of integrin-ligand interactions. *J Cell Sci* 97:585–592, 1990
5. Akiyama SK, Yamada KM: The interaction of plasma fibronectin with fibroblastic cells in suspension. *J Biol Chem* 260:4492–4450, 1985
6. Pytela R, Pierschbacher MD, Ruoslahti E: Identification and isolation of a 140kd cell surface glycoprotein with properties expected of a fibronectin receptor. *Cell* 40:191–198, 1985
7. Stepp MA, Spurr-Michaud S, Tisdale A, Elwell J, Gipson IK: $\alpha 6\beta 4$ integrin heterodimer is a component of hemidesmosomes. *Proc Natl Acad Sci USA* 87:8970–8974, 1990
8. Humphries MJ, Akiyama SK, Komoriya A, Olden K, Yamada KM: Identification of an alternatively spliced site in human plasma fibronectin that mediates cell type-specific adhesion. *J Cell Biol* 103:2637–2647, 1986
9. Humphries MJ, Komoriya A, Akiyama SK, Olden K, Yamada KM: Identification of two distinct regions of the type III connecting segment of human plasma fibronectin that promote cell type-specific adhesion. *J Biol Chem* 262:6886–6892, 1987
10. Wayner EA, Garcia-Pardo A, Humphries MJ, McDonald JA, Carter WG: Identification and characterization of the T lymphocyte adhesion receptor for an alternative cell attachment domain (CS-1) in plasma fibronectin. *J Cell Biol* 109:1321–1330, 1989
11. Mould AP, Wheldon LA, Komoriya A, Wayner EA, Yamada KM, Humphries MJ: Affinity chromatographic isolation of the adhesion receptor for the IIICS region of fibronectin and its identification as the integrin $\alpha 4\beta 1$. *J Biol Chem* 265:4020–4024, 1989
12. Komoriya A, Green LJ, Mervic M, Yamada SS, Yamada KM, Humphries MJ: The minimal essential sequence for a major cell-type specific adhesion site (CS1) within the alternatively spliced IIICS domain of fibronectin is Leu-Asp-Val. *J Biol Chem* 266:15075–15079, 1991
13. Mould AP, Komoriya A, Yamada KM, Humphries MJ: The CS5 peptide is a second site in the IIICS region of fibronectin recognized by the integrin $\alpha 4\beta 1$. Inhibition of $\alpha 4\beta 1$ function by RGD peptide homologues. *J Biol Chem* 266:3579–3585, 1991
14. Kloczewiak M, Timmons, Hawiger J: Localization of a site interacting with human platelet receptor on carboxy-terminal segment of human fibrinogen γ chain. *Biochem Biophys Res Commun* 107:181–187, 1982
15. Lam SC-T, Plow EF, Smith MA, Andrieux A, Ryckwaert J-J, Marguerie G, Ginsberg MH: Evidence that arginyl-glycyl-aspartate peptides and fibrinogen γ chain peptides share a common binding site on platelets. *J Biol Chem* 262:947–950, 1987
16. D'Souza SE, Ginsberg MH, Burke TA, Lam SC-T, Plow EF: Localization of an Arg-Gly-Asp recognition site within an integrin adhesion receptor. *Science* 242:91–93, 1988
17. D'Souza SE, Ginsberg MH, Burke TA, Plow EF: The ligand binding site of the platelet integrin receptor GPIIb-IIIa is proximal to the second calcium binding domain of its α subunit. *J Biol Chem* 265:3440–3446, 1990

18. SMITH JW, CHERESH DA: The Arg-Gly-Asp binding domain of the vitronectin receptor. Photoaffinity cross-linking implicates amino acid residues 61-203 of the β subunit. *J Biol Chem* 263:18726–18731, 1988
19. SMITH JW, CHERESH DA: Integrin ($\alpha v\beta_3$)-ligand interaction. Identification of a heterodimeric RGD binding site on the vitronectin receptor. *J Biol Chem* 265:2168–2172, 1990
20. LOFTUS JC, O'TOOLE TE, PLOW EF, GLASS A, FRELINGER AL, GINSBERG MH: A β3 integrin mutation abolishes ligand binding and alters divalent cation dependent conformation. *Science* 249:915–918, 1990
21. ARNAOUT MA: Structure and function of the leukocyte adhesion molecules CD11/CD18. *Blood* 75:1037–1050, 1990
22. D'SOUZA SE, GINSBERG MH, MATSUEDA GR, PLOW EF: A discrete sequence in a platelet integrin is involved in ligand recognition. *Nature* 350:66–68, 1991
23. CORBI AL, MILLER LJ, O'CONNOR K, LARSON RS, SPRINGER TA: cDNA cloning and complete primary structure of the α subunit of a leukocyte adhesion glycoprotein. *EMBO J* 6:4023–4028, 1987
24. SPRINGER TA: Adhesion receptors of the immune system. *Nature* 346:425–434, 1990
25. PARISE LV, HELGERSON SL, STEINER B, NANNIZZI L, PHILLIPS DR: Synthetic peptides derived from fibrinogen and fibronectin change the conformation of purified platelet glycoprotein IIb-IIIa. *J Biol Chem* 262:12597–12602, 1987